

# Associations Between Race/Ethnicity, Language, and Enrollment on Cancer Research Studies

Ogochukwu M. Ezeoke<sup>1,2,3, ID</sup>, Gary Brooks<sup>2, ID</sup>, Michael A. Postow<sup>3,4, ID</sup>, Shrujal Baxi<sup>3, ID</sup>, Soo Young Kim<sup>3, ID</sup>, Bharat Narang<sup>3, ID</sup>, Lisa C. Diamond<sup>\*,3,4,5, ID</sup>

<sup>1</sup>Department of Pediatrics, Ann and Robert H. Lurie Children's Hospital, Chicago, IL, USA

<sup>2</sup>College of Medicine, State University of New York Upstate Medical University, Syracuse, NY, USA

<sup>3</sup>Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA

<sup>4</sup>Weill Cornell Medical College, New York, NY, USA

<sup>5</sup>Department of Psychiatry and Behavioral Sciences, Memorial Sloan Kettering Cancer Center, New York, NY, USA

\*Corresponding author: Lisa C. Diamond, MD, MPH, Assistant Attending Physician, Department of Medicine, Memorial Sloan Kettering Cancer Center, 485 Lexington Avenue, 2nd Fl, New York, NY 10017, USA. Tel: +1 646 888 4246; Email: [diamondl@mskcc.org](mailto:diamondl@mskcc.org)

## Abstract

**Purpose:** The objective of this study was to determine whether differences in patients' race/ethnicity, preferred language, and other factors were associated with patient enrollment in oncology research studies.

**Patients and Methods:** We conducted a retrospective cross-sectional analysis of all adults (>18 and ≤90) seen at a large, metropolitan cancer center from 2005 to 2015, examining if enrollment to a research study, varied by race/ethnicity, preferred language, comorbidities, gender, and age.

**Results:** A total of 233 604 patients were available for initial analysis. Of these, 93 278 (39.9%) were enrolled in a research protocol (therapeutic and non-therapeutic studies). Patients who self-reported their race/ethnicity as Native, Other, Unknown, or Refuse to Answer were less likely to be enrolled on a study. Patients with one or more comorbidities, and those whose preferred language was English, were more likely to be enrolled on a research study. A logistic regression model showed that, although Non-Hispanic Black patients were more likely to have one or more comorbidities and had a higher proportion of their subset selecting English as their preferred language, they were less likely to be enrolled on a study, than our largest population, Non-Hispanic/White patients.

**Conclusions:** We identified differences in research study enrollment based on preferred language, and within race/ethnicity categories including Native-Populations, Other, Unknown or Refuse to Answer compared to Non-Hispanic/White patients. We also highlighted the lower odds of enrollment among Non-Hispanic/Black patients, in the setting of factors such as comorbidities and English language preference, which were otherwise found to be positive predictors of enrollment. Further investigation is needed to design targeted interventions to reduce disparities in oncology research study enrollment, with particular focus on language diversity.

**Key words:** healthcare disparities; language barriers; language equity; clinical trials; cancer research studies

## Implications for Practice

Diverse enrollment on cancer research studies may be impeded by language barriers and the lack of adequate resources to support a multi-lingual trial population. During trial development, this is an area where particular focus may be dedicated to increase diversity in research studies. This study highlights the impact of language on enrollment on cancer research studies, discusses the implications of language diversity on increasing multi-ethnic study enrollment, and provides methods to support multi-ethnic/multi-lingual enrollment.

## Introduction

Non-White populations, including multi-ethnic, multi-lingual populations, are more likely to present with tumors with more aggressive biological behavior compared with Non-Hispanic/White patients at initial diagnosis.<sup>1–8</sup> These factors contribute to an overall worse prognosis and decreased overall survival in these populations. Previous studies have examined the reasons for these disparities in presentation and outcome.<sup>1,2</sup> These have included assessments of differences in disease manifestation, analyses of socioeconomic and education levels, and

studies of genomic variations across race/ethnicity.<sup>2,3</sup> Several investigations indicate a higher incidence among Non-White race/ethnicity groups for certain cancers.<sup>1</sup> To understand the biology of the patients with cancer in the US, diverse cohorts are necessary; they provide clarity on any variations in cancer pathology by race/ethnicity or geographical ancestry, and promote patient centered therapy development.<sup>9–11</sup>

Literature examining overall participation in cancer research studies and investigating the role of race/ethnicity and language is sparse. Particularly regarding enrollment in

studies for rare cancers and studies involving social influences of health. Research diversity improves the generalizability of cancer therapy, improves patient-reported outcomes, and allows the development of trials that support a variety of patient needs.<sup>12-14</sup> Understanding the factors that may be associated with enrollment of study volunteers will inform the development of strategies to increase participation in cancer research studies. We aim to use the experience of a large urban cancer center to investigate how race/ethnicity, the presence of comorbidities, and preferred language may affect participation in oncology research studies.

## Methods

We performed a retrospective cross-sectional analysis of adult patients seen at an NCI-designated comprehensive cancer center between 2005 and 2015. We assessed variability in enrollment—defined as registration into any research study—based on race/ethnicity and preferred language. We collected data on patient gender and comorbidities, as these variables are part of inclusion/exclusion criteria for some research studies.

Race/ethnicity is self-reported by patients upon initial registration. Patients may select a specific race category, of the following: Asian, Black/African-American, Native-American or Native-Hawaiian, Other, Unknown, Refused to Answer or White. A free text response for Race was not an option during our study period. An ethnicity category also allows the following: Hispanic/Latinx, Not Hispanic, and Unknown. The institution in the last 10 years made the change to separate the race/ethnicity selections into 2 distinct options. In order to assess data similarly prior to and following this change, we combined both variables. If participants indicated that they were Hispanic or Latinx, this was coded as their race/ethnicity variable, otherwise only the selected Race category was recorded. Eight mutually exclusive race/ethnicity categories were established: Non-Hispanic/White (NHW), Non-Hispanic/Black (NHB), Asian (AN), Hispanic/Latinx (HL), Native Populations (NP), Other (OT), Refused to Answer (RF), and Unknown (UNK). Preferred language is self-selected by patients upon registration out of 94 options. Preferred language was dichotomized into English or Non-English. Protocol exclusion by preferred language was defined by requirements for native English or equivalent fluency. Patient gender is self-selected at time of hospital registration, out of gender options: male, female, and the option to leave blank. To avoid theorizing upon missing data on gender, we did not include patients for whom these data were missing, in final analyses.

Comorbidities were identified using ICD-10 coding in patient electronic medical records. We included diabetes, coronary artery disease and hypertension, based on their high prevalence in the adult US population.<sup>15</sup> We identified 2155 research protocols available over the time span of our study, including study types of Therapeutic ( $n = 1514$ ), Diagnostic ( $n = 265$ ), Clinical Genetic ( $n = 23$ ), Epidemiologic ( $n = 58$ ), Prevention ( $n = 19$ ), Psychosocial ( $n = 126$ ), Quality of Life ( $n = 106$ ), and Specimen Banking ( $n = 44$ ). We reviewed available eligibility criteria for protocols which had electronic eligibility documentation within the hospital record ( $n = 1647$ ), to evaluate for exclusion based on our 3 selected comorbidities. We coded exclusion criteria in studies using criteria listed on [www.clinicaltrials.gov](http://www.clinicaltrials.gov) by disease system (Cardiovascular, Endocrine, Hepatorenal, etc.), and focused on Cardiovascular and Endocrine to capture our comorbidities of interest

(diabetes, coronary artery disease, and hypertension). The presence of comorbidities was determined as positive ICD-10 code for any of the 3 selected comorbidities.

We assessed for variation in enrollment, while accounting for study exclusion criteria related to comorbidities and Non-English preference. Using unique anonymizing identifiers, patients were stratified into enrolled vs. not enrolled in at least one research study. Research studies included the following protocol types: clinical genetics, diagnostic, epidemiologic, pediatric, preventive, psychosocial, quality of life, specimen banking, and therapeutic/clinical trials. We excluded from our analysis all patients <18 and  $\geq 90$  years of age, as well as patients with no value listed for race/ethnicity, language or gender.

Descriptive analyses used frequencies (%) and means (SD) as appropriate to the data. Bivariate associations of enrollment status by variables of interest used chi-square analyses. We assessed, with NHW—the largest cohort—as the reference group, the odds ratio of enrollment on a research study by race/ethnicity, language, gender, and presence of comorbidities. We performed all analyses using SAS, version 9.4. These analyses were approved by the cancer center's IRB.

## Results

A total of 233 604 patients were available for initial review. This included all patients seen at the cancer center between 2005-2015 with a cancer diagnosis. Of these, 93 278 (39.9%) were enrolled in at least one study. [Table 1](#) shows research study enrollment by patient demographics and comorbidities. The race/ethnicity breakdown of the total initial population was: NHW: 74.6%, NHB: 5.5%, AN: 4.6%, HL: 4.2%, NP: 0.1%, OT: 0.5%, RF: 3.5%, and UNK: 3.6%. When assessing for variability in enrollment proportions by race/ethnicity alone, differences in the NHW (43.2%), NHB (44.0), AN (42.9), and HL (42.3), groups were small, although statistically significant ( $P$ -value <.0001). When including the percent enrolled within the categories of NP (27.2%), OT (38.1%), RF (33.2%), and UNK (3.4%), we see notable absolute value differences ( $P$ -value <.0001). Populations of patients selecting UNK (8445; 3.6%) or RF (8174; 3.5%), represented the fifth and sixth largest cohorts seen at the cancer center, respectively, with the NHW category being the largest (174 284; 74.6%).

A larger proportion of patients overall were of female identifying (53.2% vs 46.8%,  $P$ -value .0036). Differences in enrollment by gender were small, with female patients less likely to be enrolled compared to male patients (39.7% vs 40.2%;  $P$ -value <.0036). Although these results show statistical significance, it is difficult to determine the clinical significance of this difference. A higher proportion of patients with  $\geq 1$  comorbidities was enrolled in a study (49.0%) compared to patients with no comorbidities (33.0%) ( $P$ -value <.0001). A higher proportion, of those with English language preference, was enrolled in a study (44.2%) compared with those whose language preference was not English (31.1%) ( $P$ -value <.0001).

[Table 2](#) is a breakdown of the total initial population by race/ethnicity, by the presence of comorbidities, selection of female gender, and preferred language. The subset of patients selecting NHB had a larger proportion with  $\geq 1$  comorbidity compared with all other race/ethnicity categories (58.4% vs NHW: 45.6%, AN: 43.2%, HL: 48.0%, NP: 50.6%, OT:

**Table 1.** Research study enrollment by race/ethnicity, gender, comorbidity status, and English language preference.

	All subjects		Enrolled		Not enrolled		P-value <sup>a/**</sup>
	N	Percent of total	N	Group percent	N	Group percent	
Race/ethnicity categories							
Non-Hispanic/White	174 286	74.6	75 304	43.2	98 982	56.8	
Non-Hispanic/Black	12 963	5.5	5702	44.0	7261	56.0	
Asian	10 805	4.6	4639	42.9	6166	57.1	
Hispanic/Latinx	9727	4.2	4116	42.3	5611	57.7	
Native populations	162	0.1	44	27.2	118	72.8	<.0001 *
Other	1119	0.5	426	38.1	693	61.9	
Refused to answer	8174	3.5	2714	33.2	5460	66.8	
Unknown	8445	3.6	290	3.4	8155	96.6	
Missing <sup>+</sup>	7923	3.4	43	0.5	7880	99.5	
Gender							
Female	124 345	53.2	49 307	39.7	75 038	60.3	.0036 *
Male	109 254	46.8	43 969	40.2	65 285	59.8	
Unknown (missing <sup>+</sup> )	5	0.0	2	40.0	3	60.0	
Comorbidities							
1 or more	100 737	43.1	49 398	49.0	51 339	51.0	
None	132 867	56.9	43 880	33.0	88 987	67.0	<.0001 *
English preferred							
Yes	185 587	79.4	81 984	44.2	103 603	55.8	
No	11 407	4.9	3544	31.1	7863	68.9	<.0001 *
Missing <sup>+</sup>	36 610	15.7	7750	21.2	28 860	78.8	
Mean age (years) (±SD)	57.9(±16.0)	n/a	57.7(±13.7)	n/a	58.9(±14.6)	n/a	<.0001 **
Total	233 604	100.0	93 278	39.9	140 326	60.1	

<sup>a</sup>P-values are from Chi-square tests of independence.<sup>\*\*</sup>P-value from Mann-Whitney U test.<sup>+</sup>Patients with missing data were not included in final analysis**Table 2.** Total population: comorbidities, gender, and English language preference by race/ethnicity categories.

	Has 1 or more comorbidities			Female gender		English language preferred	
	Total population	Total <sup>a</sup>	% of total population	Total <sup>b</sup>	% of total population	Total <sup>c</sup>	% of total population
Race/ethnicity categories							
Non-Hispanic/White	174 286	79 546	(45.6)	91 281	(52.4)	148 253	(85.1)
Non-Hispanic/Black	12 963	7,569	(58.4)	7,873	(60.7)	11 573	(89.3)
Asian	10 805	4666	(43.2)	6393	(59.2)	7899	(73.1)
Hispanic/Latinx	9727	4671	(48.0)	5682	(58.4)	6416	(66.0)
Native populations	162	82	(50.6)	89	(54.9)	138	(85.2)
Other	1119	459	(41.0)	571	(51.0)	829	(74.1)
Refused to answer	8174	3175	(38.8)	4039	(49.4)	6934	(84.8)
Unknown	8445	496	(5.9)	4241	(50.2)	1729	(20.5)
Total	233 604	100 737	(43.1)	124 345	(53.2)	185 587	(79.4)

All values are N (%).

All P-values &lt;.0001 (Chi-square test of independence).

<sup>a</sup>Total with 1 or more comorbidities.<sup>b</sup>Total who identified as female at hospital registration.<sup>c</sup>Total who selected English as the clinical language preference.

41.0%, RF: 38.8%, and UNK: 5.9%; P-value <.0001). The NHB population was most likely to report English as their preferred language, while patients in the UNK category,

represented the smallest proportion of patients to do so (NHB: 89.3% vs NHW: 85.1%, AN: 73.1%, HL: 66.0%, NP: 85.2%, OT: 74.1%, RF: 84.8%, and UNK: 20.5%; P-value <.0001).

**Table 3.** Total population: subset of patients with no comorbidities by English language preference.

	English language preference			Non-English language preference		
	Total*	Total* enrolled	% Total* enrolled	Total^	Total^ enrolled	% Total^ enrolled
Race/ethnicity categories						
Non-Hispanic/White	80 192	31 986	(39.9)	1529	398	(26.0)
Non-Hispanic/Black	4712	1831	(38.9)	44	6	(13.6)
Asian	4639	1967	(42.4)	929	282	(11.7)
Hispanic/Latinx	3577	1545	(43.2)	899	268	(13.6)
Native populations	66	16	(24.2)	4	0	(0.0)
Other	478	182	(38.1)	71	20	(28.2)
Refused to answer	4264	1299	(30.5)	628	119	(18.9)
Unknown	1403	84	(5.9)	31	4	(12.9)
Totals	101 097	38 926	(38.5)	4230	1098	(25.9)

All values are N (%).

All *P*-values <.0001 (Chi-square test of independence).

\*No comorbidities with English language preference.

^No comorbidities with Non-English language preference.

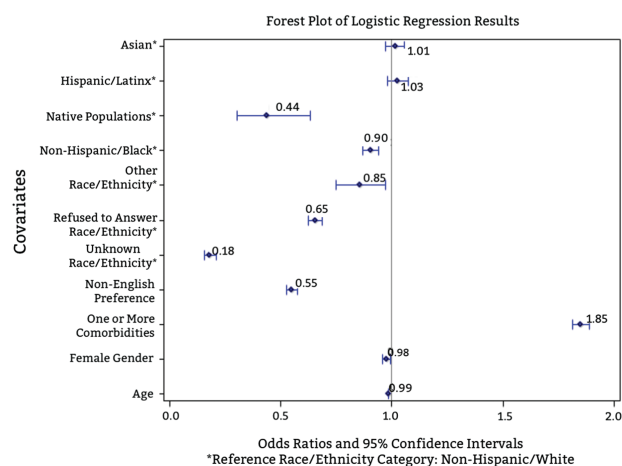
Table 3 examines patients without comorbidities, highlighting the variation by race/ethnicity between those selecting English, and those selecting any of the 94 other languages. Of patients without comorbidities, only 5.9% of those who selected both UNK and English, are enrolled. This is the lowest overall proportion when compared with the percent enrolled within all other race/ethnicity categories (5.9% vs NHW: 39.9%, NHB: 38.9%, AN: 42.4%, HL: 43.2%, NP: 24.2%, OT: 38.1%, and RF: 30.5%; *P*-value <.0001). Among those who preferred to communicate in a language other than English, the NP and AN categories were the smallest proportion of those enrolled in a study (NP: 0% and AN: 11.7% vs NHW: 26.0%, NHB: 13.6%, HL: 13.6%, OT: 28.2%, RF: 18.9%, UNK: 12.9%; *P*-value <.0001).

Of the 1647 studies with available eligibility criteria, 1451 (88.1%) excluded patients based on cardiovascular or endocrine disease, which include our selected comorbidities. Of these studies, 218 (13.2%) excluded patients based on non-English preferred language. Exclusion by preferred language was defined as English fluency insufficient to the point of requiring study translation.

Figure 1 shows logistic regression assessing multifactorial odds ratio of enrollment. There were reduced odds of enrollment in the NHB (OR: 0.91, 95% CI: 0.87, 0.94), RF (OR: 0.65, 95% CI: 0.62, 0.69), NP (OR: 0.44, 95% CI: 0.31, 0.63), UNK (OR: 0.18, 95% CI: 0.16, 0.21), and OT (OR: 0.85, 95% CI: 0.75, 0.97) categories. Patients who preferred non-English languages had reduced odds of enrollment in a research study (OR: 0.55, 95% CI: 0.53, 0.58). Logistic regression showed slightly reduced odds for female identifying patients being enrolled (OR = 0.98, 95% CI: 0.96, 0.99). Patients with one or more comorbidities had increased odds of enrollment (OR: 1.85, 95% CI: 1.81, 1.89). For each 1-year increase in age, the odds of being enrolled in a study decreases by 0.015.

## Discussion

Our data show that statistically significant differences exist in enrollment by race/ethnicity, on a wide range of research studies. The practical significance of differences seen between

**Figure 1.** Logistic regression for research study enrollment.

NHW, NHB, HL, and AN is difficult to determine, with the percent enrollment, per group, all just above 40%. On the contrary, there is a potentially actionable difference in the enrollment of those in the NP, OT, RF, and UNK groups, with a range from 3.4% in UNK to 38.1% in OT. Although all results show statistical significance, with *P*-values <.05, we must apply nuance and recognize that further investigation is needed to scrutinize and interpret small differences seen between NHW, NHB, HL, and AN groups. The sharper disparity seen both between NP, OT, RF, and UNK, and as compared to NHW, NHB, HL, and AN, warrants examination for action, particularly as relates to patients selecting race/ethnicity as UNK. This group had the lowest proportion selecting English as the preferred language for clinical care and the lowest proportion with ICD-10 codes for any of our selected comorbidities.

We see that the presence of comorbidities appears to be a positive factor toward enrollment in a research study. The NHB population had higher incidences of one or more comorbidities, however they still appear to have lower odds of being enrolled overall on a study, as seen in our logistic regression. Additionally, although the selection of English as a preferred



language appears to be associated with an increased odds of enrollment on a research protocol, this, once more, does not hold true within our NHB population, despite having a high percentage of patients with English as their preferred language. While there is an association between language selection and enrollment on a research study, there remains nuance in how these results can be interpreted across different populations.

As we uncover the role of language on enrollment, we recognize that studies have highlighted difficulties in research informed consent, when language barriers exist.<sup>16-18</sup> While we observed exclusion in studies based on preferred language, this may be more clearly explained as exclusion based on challenges in translating study materials into non-English languages and gaps in resources to devote to professional interpreter use in study visits. The ability to translate into plain English the implications and requirements of a research study is one which oncologists practice daily. However, the insertion of a third party into the discussion, such as an interpreter, the lack of an established relationship with a physician, and the absence of trust in the healthcare system could all limit successful communication.<sup>16,19</sup> Regarding the protocol documents, including consent and any additional patient forms associated with the study, there is a dependence on the industry sponsor or NIH-fund to facilitate the translation to any potential written language for use by the patient. However, even with translation of documents, the consent process requires the additional layer of culturally informed clinical discussions in the appropriate language.

It appears that language barriers may not be the cause of lower cancer research enrollment in our NHB population. The unfortunate historical context of the institution of medicine and the NHB community includes numerous actions which have significantly impeded trust. These actions include prominent examples of large-scale NHB research enrollment include the Tuskegee Syphilis Study and the Baltimore Lead Paint Study.<sup>20-22</sup> The second lowest proportion for enrollment is seen in NP, who, like the NHB population, has a high ratio of the cohort selecting English as their primary clinical communication language, and has a high incidence of one or more comorbidities. Here, once more we evaluate these results by looking at historical context. In reviewing the literature, Findling et al show reports of discrimination by Native Americans, and potential avoidance of healthcare interaction—which would be required for participation in a research study—for fear of discrimination. This may be a consequence of historical injustices—as discussed by Walters and Simoni—that has tainted the trust between Native Americans and the healthcare system.<sup>23-25</sup>

Medicine will require rebuilding of trust, educating clinicians about the context with which NP, NHB and limited English proficient (LEP) patients arrive at the hospital. Through this exercise we may be able to better address the communication barriers that exist within cancer trial informed consent. The cancer center is working to increase research enrollment for previously underrepresented participants, including LEP patients, through a core facility which advises principal investigators on community outreach and engagement.

It is worth noting that patients with the presence of at least one comorbidity represented a larger proportion of total enrolled patients. This is perhaps secondary to an established relationship with a healthcare provider, and potentially

positive experiences. In examining the literature, it is clear that having a relationship within oncology or clinical care plays a significant role in commitment of a patient to joining a clinical trial.<sup>26</sup> Enrollment in research studies is often guided by the presence of, and ability to manage comorbidities in patients, particularly as there is a perceived increase in toxicity in these trial participants.<sup>27,28</sup> These stringent requirements do not necessarily extend to the enrollment of patients on non-therapeutic research protocols. There were 2155 research protocols available over the time span of our study, of which therapeutic ( $n = 1514$ ) and diagnostic ( $n = 265$ ) studies made up 82.6%. Within this selection, only 76.4% ( $n = 1647$ ) had available inclusion and exclusion criteria in an accessible electronic format. Additional included study types were Clinical Genetic ( $n = 23$ ), Epidemiologic ( $n = 58$ ), Prevention ( $n = 19$ ), Psychosocial ( $n = 126$ ), Quality of Life ( $n = 106$ ), and Specimen Banking ( $n = 44$ ). Within individual trials, we must acknowledge that there are likely specific exclusion criteria which our broad based evaluation would not capture, particularly regarding clinical exclusions or comorbidities. Additionally, the identification of race/ethnicity differences in enrollment within therapeutic versus non-therapeutic trials would have added significant value to our study; it would have allowed us to evaluate the role of therapeutic benefit in patient enrollment decisions. As many patients were enrolled in more than one study over our assessment period, we focused on enrollment to any study rather than enrollment by trial type, to avoid duplicate counting. The increased enrollment among patients with comorbidities was not observed in the NP or NHB subgroups. There are a variety of possible reasons for this finding, including mistrust in medical systems as noted above. This mistrust among NP and NHB patients, with comorbidities, may be exacerbated by increased exposure to the medical system, where reports of discrimination are seen at greater proportions when compared with NHW populations.<sup>23-25,29</sup> We must also consider that clinicians may be reluctant to refer patients to clinical trials, due to reasons including lack of resources to address minoritized populations such as NP and NHB, with culturally appropriate recruitment and retention strategies.<sup>30,31</sup> The cancer center has responded to these known biases through its Office of Health Equity. Specific strategies currently being used include a detailed evaluation of disparities in clinical trial enrollment for any open studies at the center, an assessment of possible barriers to diverse enrollment for studies in the development phase and the identification of trial enrollment benchmarks to strive toward.

Our study underlines stark differences in enrollment in the categories of NP (27.2%), OT (38.1%), RF (33.2%), and UNK (3.4%) compared with other categories and shows that those who selected UNK were least likely to have selected English as their preferred language. The reporting of race/ethnicity is part of the registration process of new patients to the cancer center, to allow the evaluation of equitable care throughout the course of a patient's treatment.<sup>32</sup> It is a voluntary segment of the registration process and our results compel us to consider 2 factors: the role of race/ethnicity in perceived medical care and the understanding of US-based social constructions of race/ethnicity. The perception that selecting one race/ethnicity category may impact the quality of care being delivered is not unique amongst, for instance, the NHB community.<sup>33</sup> A patient may perhaps perceive that selecting UNK may be preferable when hoping for the excellent care that is

offered at a large institution such as the study cancer center. The second consideration we have is the understanding of these categories within the international, non-US population. The social construct of race remains a concept that does not fully capture the geographical ancestry of patients across the globe who may present to a large center for cancer care. Patients may select UNK if they do not find the race/ethnicity categories applicable to their own racial/ethnic identity. The selection of UNK, combined with any additional language barriers, may also create a perception among clinicians that a language barrier and informed consent might be challenging.<sup>34</sup> The result, regardless, is a decrease in enrollment of patients who do not fit within the demarcations of historical race/ethnicity categories.

Previous research that has highlighted disparities in enrollment in cancer research studies, has focused on specific cancer diagnoses.<sup>3,35-39</sup> Our analysis is unique in that we evaluate all available studies, including tumor banking studies, which may not, for instance, require the same time or trust commitment as might therapeutic trials.<sup>40</sup> Cancer research studies allow us to tackle a variety of questions related to patient care, including among many, quality of life, treatment response, and cancer genomics. Given the broad nature of our study, we are able to address the question of enrollment influences across the spectrum of research studies, with and without the consideration of direct therapeutic benefit to the enrolled patient. Within the selection of protocols which were available during our study period, the majority were interventional in nature including therapeutic ( $n = 1514$ , 70.3%) and diagnostic clinical trials ( $n = 265$ , 12.3%); however, several other study types were available. Our approach also provides a reflection on additional broad variables which may impact enrollment, including pregnancy as an exclusion criteria for many studies,<sup>41</sup> and the availability of studies for disease affecting one gender over the other. For instance, an analysis of NIH-funded studies by Mirin et al.,<sup>42</sup> shows increased overall funding for studies affecting men whereas there was on average underfunding for studies targeted at diseases primarily affecting women.<sup>41</sup> These factors could certainly impact the enrollment numbers which we found, particularly with respect to gender where women (39.7%) were enrolled at a lower proportion compared to men (40.2%). We are able to take a broad approach to ameliorating the impact of perceived language barriers, as well as mistrust of the medical community. Critically, enrollment goals for every study must be clarified during trial development; ensuring a diverse study population requires stating unambiguously the need for diverse recruitment. Meaningful goals for multi-ethnic enrollment should at least reflect the population demographics. To facilitate this effort, the inclusion of community leaders and clinicians as collaborative investigators is necessary. Although there should be ongoing diversification of the primary site clinician workforce, seeking out local practitioners from diverse, multi-lingual communities may ease discussions around research enrollment with patients who may be initially hesitant to be involved. Studies have repeatedly highlighted the improvement in physician-patient communication, where race concordance exists,<sup>43</sup> and this data should be used to the advantage of research study developers. Finally in reviewing the criteria for enrollment on research studies, ensuring diversity amongst reviewers, may allow nuanced conversations around who may be unintentionally excluded, via criteria which may not significantly impact the success of the study.

## Limitations

Although this is a large metropolitan hospital which serves 23 counties of an estimated 18 million people, with 53% identifying as Non-White, a community health needs assessment conducted by the cancer center, found that in 2015, <20% of the cancer center's patient population was diverse. The assessment reported perceptions, by patients, of exclusivity.<sup>44,45</sup> We acknowledge the difficulty in capturing a diverse cohort of patients with this perception likely influencing patient presentation to the hospital.

An additional limitation has been our ability to fully capture the role of comorbidities; as clinician assessment of these comorbidities was often the determining factor for participation on a trial, this data would have been beneficial for analysis. Comorbidity data in the electronic medical record are often incomplete<sup>46</sup> and the center EMR was unfortunately not an exception, with sometimes inconsistent inputs of each clinical diagnosis, recorded as ICD-10 codes. Regarding which diagnoses to include in our study analysis, our selection was thus based on most common US adult chronic health conditions, which were also most consistently recorded in the medical record. We recognize the limitation of this list, not including a comprehensive selection of all comorbidities which may impact patient enrollment. We also acknowledge the limitation within the medical record surrounding timing of comorbidity diagnosis; as these records are timed based on clinician input into the EMR, they do not always correspond to initial diagnosis, and do not consistently reflect pre-enrollment comorbidities.

We also recognize that the availability of clinical trials varies by cancer type and that there are variations in cancer incidence by patient geographical ancestry. Further analysis, stratifying disease type by our race/ethnicity categories would have clarified these variations. We also acknowledge the limitation of the race/ethnicity categories, as they have been changed in the last decade. The ability to assess race/ethnicity variation in healthcare delivery, and thus identify opportunities for improving equitable care is limited when patients may not feel comfortable selecting a racial category. Equity research is improved when patients are provided with the options to most accurately describe their identity. Although we have found decreased enrollment amongst patients in the Native-Populations, Other, Unknown or Refuse to Answer categories, these findings would be more specifically actionable with the inclusion of more personalized patient data. We are able to infer potential barriers, based on the literature on clinical trial enrollment<sup>47-49</sup>, but it is a challenge to investigate more precisely into variations by race/ethnicity without improvements in how race/ethnicity data is captured. We also acknowledge the limitation of a binary-favoring gender system that does not reflect a diverse gender identity. The cancer center has responded to these known biases through its Office of Health Equity with a re-evaluation of the process for demographic data capture, including how race/ethnicity and gender information are collected.

## Conclusions

Our study highlights potential areas for improvement in the approach to increasing diversity in cancer research enrollment; while we continue to work toward increasing racial/ethnic diversity in cancer research enrollment to reflect the

patient population who would benefit from research studies, we must realize that language barriers may also be impeding our efforts. We must also acknowledge the potential positive impact of culturally considered care on recruitment to cancer research studies should not be ignored. Enrollment in cancer research studies by a diverse cohort allows adequate interpretation of results that is beneficial to a multi-national patient population and provides insight into potential differences in pathology, as well as drug pharmacokinetics and pharmacodynamics across geographical ancestry.<sup>2,50</sup> Language barriers create limited research study populations, in addition to known disparities in healthcare.

## Acknowledgments

The authors would like to thank the ASCO Conquer Cancer Foundation and the AMA Seed Grant Foundation for their funding of this research.

## Funding

The research and data analysis leading to these results was funded by the American Society of Clinical Oncology Conquer Cancer Foundation (Grant ID: 10372) and the American Medical Association Seed Grant Foundation.

## Conflict of Interest

**Shrujal Baxi:** Flatiron Health, a Roche subsidiary (E [former], OI), Iterative Scopes (E); **Michael A. Postow:** Array BioPharma, Aduro, BMS, Incyte, Merck, NewLink Genetics, Novartis, Eisai, Pfizer, Chugai (SAB), RGenix, Infinity, BMS, Merck, Array BioPharma, Novartis (RF [institutional]). The other authors indicated no financial relationships.

(C/A) Consulting/advisory relationship; (RF) Research funding; (E) Employment; (ET) Expert testimony; (H) Honoraria received; (OI) Ownership interests; (IP) Intellectual property rights/inventor/patent holder; (SAB) Scientific advisory board.

## Author Contributions

Conception/design: O.M.E., M.A.P., L.C.D. Provision of study material or patients: O.M.E., M.A.P., L.C.D. Collection and/or assembly of data: O.M.E., S.Y.K., B.N. Data analysis and interpretation: G.B., S.B., S.Y.K., B.N., L.C.D. Manuscript writing: All authors. Final approval of manuscript: All authors.

## Data Availability

The data underlying this article will be shared on reasonable request to the corresponding author.

## References

- Howlader N, Noone AM, Krapcho M, et al. *SEER Cancer Statistics Review, 1975-2012*. National Cancer Institute; April 2015. [http://seer.cancer.gov/csr/1975\\_2012/](http://seer.cancer.gov/csr/1975_2012/), based on November 2014 SEER data submission, posted to the SEER web site.
- Trevino M, Padalecki S, Karnad A, et al. The development of a minority recruitment plan for cancer clinical trials. *J Community Med Health Educ*. 2013;3(5):1000230. <https://doi.org/10.4172/2161-0711.1000230>.
- Brown SD, Partee PN, Feng J, et al. Outreach to diversify clinical trial participation: a randomized recruitment study. *Clin Trials*. 2015;12(3):205-211. <https://doi.org/10.1177/1740774514568125>.
- Zullig LL, Fortune-britt AG, Rao S, et al. Enrollment and racial disparities in cancer treatment clinical trials in North Carolina. *N C Med J*. 2016;77(1):52-58. <https://doi.org/10.18043/ncm.77.1.52>.
- Stubblefield J, Kelly B. Melanoma in non-Caucasian populations. *Surg Clin N Am*. 2014;94:1115-1126.
- Park SL, Marchand LL, Wilkens LR, et al. Risk factors for malignant melanoma in White and non-White/non-African American populations: the multiethnic cohort. *Cancer Prev Res (Phila)*. 2012;5:423-434.
- Cress RD, Holly EA. Incidence of cutaneous melanoma among non-Hispanic whites, Hispanics, Asians, and blacks: an analysis of California cancer registry data, 1988-93. *Cancer Causes and Control* 1997;5:246-252.
- Hu S, Parmet Y, Allen G, et al. Disparity in melanoma: a trend analysis of melanoma incidence and stage at diagnosis among whites, Hispanics, and blacks in Florida. *Arch Dermatol*. 2009;145:1369-1374. <https://doi.org/10.1001/archdermatol.2009.302>.
- Ahaghotu C, Tyler R, Sartor O. African American participation in oncology clinical trials-focus on prostate cancer: implications, barriers, and potential solutions. *Clin Genitourin Cancer*. 2016;14(2):105-116. <https://doi.org/10.1016/j.clgc.2015.12.003>.
- Aristizabal P, Singer J, Cooper R, et al. Participation in pediatric oncology research protocols: racial/ethnic, language and age-based disparities. *Pediatr Blood Cancer*. 2015;62(8):1337-1344. <https://doi.org/10.1002/pbc.25472>.
- Leiter A, Diefenbach MA, Doucette J, Oh WK, Galsky MD. Clinical trial awareness: Changes over time and sociodemographic disparities. *Clin Trials*. 2015;12(3):215-223. <https://doi.org/10.1177/1740774515571917>.
- Haynes-Maslow L, Godley P, Dimartino L, et al. African American women's perceptions of cancer clinical trials. *Cancer Med* 2014;3(5):1430-1439. <https://doi.org/10.1002/cam4.284>.
- Simon MA, De La Riva EE, Bergan R, et al. Improving diversity in cancer research trials: the story of the cancer disparities research network. *J Cancer Educ*. 2014;29(2):366-374. <https://doi.org/10.1007/s13187-014-0617-y>.
- Byrne MM, Tannenbaum SL, Glück S, Hurley J, Antoni M. Participation in cancer clinical trials: why are patients not participating?. *Med Decis Making*. 2014;34(1):116-126. <https://doi.org/10.1177/0272989X13497264>.
- Chronic Diseases in America | CDC. 2019. Retrieved 2019, from <https://www.cdc.gov/chronicdisease/resources/infographic/chronic-diseases.htm>
- Brelsford KM, Ruiz E, Hammack CM, Beskow LM. Improving translation and cultural appropriateness of Spanish-language consent materials for biobanks. *Ethics Hum Res*. 2019;41(5):16-27. <https://doi.org/10.1002/eahr.500028>.
- Hanrahan D, Sexton P, Hui K, et al. Linguistic and cultural challenges in communication and translation in US-Sponsored HIV prevention research in emerging economies. *PLoS One*. 2015;10(7):e0133394. <https://doi.org/10.1371/journal.pone.0133394>.
- Lee JS, Pérez-Stable EJ, Gregorich SE, et al. Increased access to professional interpreters in the hospital improves informed consent for patients with limited English proficiency. *J Gen Intern Med*. 2017;32(8):863-870. <https://doi.org/10.1007/s11606-017-3983-4>.
- Altman MR, Oseguera T, McLemore MR, Kantrowitz-Gordon I, Franck LS, Lyndon A. Information and power: Women of color's experiences interacting with health care providers in pregnancy and birth. *Soc Sci Med*. 2019;238(1982):112491. <https://doi.org/10.1016/j.socscimed.2019.112491>.
- Kaiser, J. Human subjects. court rebukes hopkins for lead paint study. *Science*. 293(5535):1567-69. <https://doi.org/10.1126/science.293.5535.1567b>. American Association for the Advancement of Science.
- McCallum JM, Arekere DM, Green BL, Katz RV, Rivers BM. Awareness and knowledge of the U.S. public health service syphilis



- study at Tuskegee: implications for biomedical research. *J Health Care Poor Underserved*. Nov. 2006;17(4):716-733. <https://doi.org/10.1353/hpu.2006.0130>.
22. Scharff DP, Mathews KJ, Jackson P, et al. More than Tuskegee: understanding mistrust about research participation. *J Health Care Poor Underserved*. 2010;21(3):879-897. <https://doi.org/10.1353/hpu.0.0323>.
  23. Findling MG, Casey LS, Fryberg SA, et al. Discrimination in the United States: experiences of native Americans. *Health Serv Res*. 2019;54(Suppl 2):1431-1441. <https://doi.org/10.1111/1475-6773.13224>.
  24. Behring M, Hale K, Ozaydin B, et al. Inclusiveness and ethical considerations for observational, translational, and clinical cancer health disparity research. *Cancer* 2019;125(24):4452-4461. <https://doi.org/10.1002/cncr.32495>.
  25. Walters KL, Simoni JM. Reconceptualizing native women's health: an "indigenist" stress-coping model. *Am J Public Health*. 2002;92(4):520-524. <https://doi.org/10.2105/ajph.92.4.520>.
  26. Comis, Robert L., Miller Jon D, Colaizzi Diane D, Kimmel Linda G. Physician-related factors involved in patient decisions to enroll onto cancer clinical trials. *J Oncol Pract*. 5(2), 2009, 50-6, <https://doi.org/10.1200/JOP.0922001>. American Society of Clinical Oncology.
  27. Van Spall, Harriette G. C. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *JAMA*. 297(11):1233-40. <https://doi.org/10.1001/jama.297.11.1233>. American Medical Association.
  28. Duma, Narjust, Kothadia Sejal M, Azam Tariq U, et al. Characterization of comorbidities limiting the recruitment of patients in early phase clinical trials. *Oncologist*. 2019;24(1):96-102, <https://doi.org/10.1634/theoncologist.2017-0687>. AlphaMed Press.
  29. Bleich SN, Findling MG, Casey LS, et al. Discrimination in the United States: experiences of Black Americans. *Health Serv Res*. 2019;54 (Suppl 2):1399-1408. <https://doi.org/10.1111/1475-6773.13220>.
  30. McCaskill-Stevens W, McKinney MM, Whitman CG, Minasian LM. Increasing minority participation in cancer clinical trials: the minority-based community clinical oncology program experience. *J Clin Oncol*. 2005;23(22):5247-5254. <https://doi.org/10.1200/JCO.2005.22.236>.
  31. Conn VS, Enriquez M, Ruppert TM, Chan KC. Cultural relevance in medication adherence interventions with underrepresented adults: systematic review and meta-analysis of outcomes. *Prev Med*. 2014;69:239-247. <https://doi.org/10.1016/j.ypmed.2014.10.021>.
  32. National Research Council. 2004. *Eliminating Health Disparities: Measurement and Data Needs*. The National Academies Press. <https://doi.org/10.17226/10979>.
  33. Lillie-Blanton, M. Race, ethnicity, and the health care system: public perceptions and experiences. *Medical Care Res Rev*. 2000;57(Suppl 1):218-35, <https://doi.org/10.1177/1077558700057001S10>. Sage Publications.
  34. Hillyer, Grace C., Beauchemin Melissa, Hershman Dawn L, et al. Discordant attitudes and beliefs about cancer clinical trial participation between physicians, research staff, and cancer patients. *Clin Trials*. Apr. 2020;17(2):184-194 <https://doi.org/10.1177/1740774520901514>. Arnold.
  35. Scalici J, Finan MA, Black J, et al. Minority participation in Gynecologic Oncology Group (GOG) studies. *Gynecol Oncol*. 2015;138(2):441-444. <https://doi.org/10.1016/j.ygyno.2015.05.014>.
  36. Trantham LC, Carpenter WR, DiMartino LD, et al. Perceptions of cancer clinical research among African American men in North Carolina. *J Natl Med Assoc*. 2015;107(1):33-41. [https://doi.org/10.1016/S0027-9684\(15\)30007-9](https://doi.org/10.1016/S0027-9684(15)30007-9).
  37. Doamekpor LA, Zuckerman DM. Lack of diversity in cancer drug clinical trials may exacerbate racial disparities in mortality rates. *Cancer Epidemiol*. 2014;38:645-646. <https://doi.org/10.1016/j.canep.2014.06.007>.
  38. Bickell NA, Wang JJ, Oluwale S, et al. Missed opportunities: racial disparities in adjuvant breast cancer treatment. *J Clin Oncol*. 2006;24(9):1357-1362. <https://doi.org/10.1200/JCO.2005.04.5799>.
  39. Hawk ET, Habermann EB, Ford JG, et al. Five NCI-designated cancer centers' data collection on racial/ ethnic minority participation in therapeutic trials – a current view and opportunities for improvement. *Cancer* 2014;120(Suppl 7):1113-1121. <https://doi.org/10.1002/cncr.28571>.
  40. Borno, Hala T, Zhang Li, Siegel Adam, Chang Emily, Ryan Charles J. At what cost to clinical trial enrollment? A retrospective study of patient travel burden in cancer clinical trials. *The Oncologist*. Oct. 2018;23(10):1242-1249, <https://doi.org/10.1634/theoncologist.2017-0628>. AlphaMed Press
  41. Committee Opinion No. 646: Ethical Considerations for Including Women as Research Participants. *Obstet Gynecol*. 2016;127(5):e100-e107. <https://doi.org/10.1097/AOG.0000000000001150>.
  42. Mirin AA. Gender disparity in the funding of diseases by the U.S. National Institutes of Health. *J Womens Health (Larchmt)*. 2021;30(7):956-963. <https://doi.org/10.1089/jwh.2020.8682>.
  43. Street RL Jr, O'Malley KJ, Cooper LA, Haidet P. Understanding concordance in patient-physician relationships: personal and ethnic dimensions of shared identity. *Ann Fam Med*. 2008;6(3):198-205. <https://doi.org/10.1370/afm.821>.
  44. MSKCC Community Service Plans. 2016. Community Health Needs Assessment, 2(1). Retrieved 2019, from [https://www.mskcc.org/sites/default/files/node/39068/documents/2016-msk-chna-results-final\\_revised.pdf](https://www.mskcc.org/sites/default/files/node/39068/documents/2016-msk-chna-results-final_revised.pdf).
  45. Furman Center Neighborhood Profile. 2017. NYU Furman Center's CoreData.nyc. Retrieved 2019, from <https://furmancenter.org/neighborhoods/view/upper-east-side>.
  46. Baier AW, Snyder DJ, Leahy IC, Patak LS, Brustowicz RM. A shared opportunity for improving electronic medical record data. *Anesth Analg*. 2017;125(3):952-957. <https://doi.org/10.1213/ANE.0000000000002134>.
  47. Heller C, Balls-Berry JE, Nery JD, et al. Strategies addressing barriers to clinical trial enrollment of underrepresented populations: a systematic review. *Contemp Clin Trials*. 2014;39(2):169-182. <https://doi.org/10.1016/j.cct.2014.08.004>.
  48. Wenzel JA, Mbah O, Xu J, et al. A model of cancer clinical trial decision-making informed by African-American cancer patients. *J Racial Ethn Heal Disparities* 2015;2(2):192-199. <https://doi.org/10.1007/s40615-014-0063-x>.
  49. Friedman DB, Foster C, Bergeron CD, Tanner A, Kim SH. A qualitative study of recruitment barriers, motivators, and community-based strategies for increasing clinical trials participation among rural and urban populations. *Am J Health Promot*. 2015;29(5):332-338. <https://doi.org/10.4278/ajhp.130514-QUAL-247>.
  50. Tan DSW, Mok TSK, Rebbeck TR. Cancer genomics: diversity and disparity across ethnicity and geography. *J Clin Oncol*. 2016;34(1):91-101. <https://doi.org/10.1200/JCO.2015.62.0096>.